

03043
lej

UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO

COLEGIO DE CIENCIAS Y HUMANIDADES

UNIDAD ACADEMICA
DE LOS
CICLOS PROFESIONAL Y DE POSGRADO

INSTITUTO DE INVESTIGACIONES
EN MATEMATICAS APLICADAS Y EN SISTEMAS
ESPECIALIZACION EN ESTADISTICA APLICADA

ENSAYO DEL DISEÑO DE MUESTRA PARA
UNA ENCUESTA EN HOGARES
SOBRE CARACTERISTICAS
SOCIO-ECONOMICAS-DEMOGRAFICAS
DE CUSMCA-ECUADOR

TRABAJO FINAL PARA OBTENER EL DIPLOMA DE
ESPECIALISTA EN ESTADISTICA APLICADA

**TESIS CON
FALLA DE ORIGEN**

Presenta:
Haycés Alvarado Vintimilla

Septiembre

1983



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

INDICE:

	Páginas
Introducción	1
1.- Descripción de la zona de estudio	4
2.- Niveles de generación de Estadísticas.	9
3. Definición de Unidades de Muestreo y Estratificación:	11
U=1 Unidades Primarias.	
U=2 Unidades Secundarias.	
U=3 Unidades de tercera y cuarta etapa.	
E=1 Estratificación.	
E=2 Estratificación del "Area Metropolitana de Cuenca".	
4. Esquema de Muestreo.	22
5. Estimadores:	26
Notación.	
Estimador de un total.	
Varianza del estimador de un Total	
Estimador de la Varianza.	
Estimador de una Media.	
Error Cuadrático Medio.	
Estimador del Error Cuadrático Medio.	
6. Conclusiones.	33
7. Definiciones.	34
8...Bibliografía.	35

INTRODUCCION:

Los motivos que me llevaron a elegir el tema; "ENSAYO DEL DISEÑO DE MUESTRA PARA UNA ENCUESTA EN HOGARES SOBRE CARACTERISTICAS SOCIO-ECONOMICO-DEMOGRAFICAS DE CUENCA-ECUADOR", como trabajo final de la "Especialización en Estadística Aplicada" son los siguientes: Dada mi profesión de Arquitecto Urbanista y las posibilidades de aplicar la Estadística en la misma, una de las materias estudiadas de mayor aplicación sería la de "Muestreo", sobre todo relacionado con la "Problemática Urbana", en este sentido el objetivo es contar con una información socio-económica-demográfica lo suficientemente rápida, adecuada y precisa para el diagnóstico, análisis y predicción de los fenómenos urbanos como el crecimiento y forma de distribución de la población en el territorio, en su doble efecto de concentración y dispersión y las modificaciones de las actividades socioeconómicas urbanas. (Población por ramas de actividad ingresos, empleo, vivienda, etc.)

Por lo tanto este trabajo tiene por objeto realizar un diseño de muestra para disponer de información rápida y a corto plazo para fines de planificación e investigación.

Los investigadores y planificadores con frecuencia necesitan respaldar sus tesis con datos estadísticos ,-

los mismos que por falta de precaución muchas veces son mal obtenidos o mal interpretados, provocando errores en las conclusiones de dichos trabajos. El Muestreo Probabilístico es una manera de poder estimar los valores de los parámetros de una población y tener una medida -- del error en que se puede incurrir al dar dicha estimación, obteniéndose por lo tanto mayor calidad a menor costo, los censos en cambio por sus características son muy complicados y requieren de grandes recursos económicos, realizándose en períodos de tiempo sumamente amplios, sirviendo entonces el Muestreo como un excelente auxiliar.

En lo que se refiere a la Planificación Urbana, es necesario mantener actualizados una serie de datos como: población, sueldos y salarios, gastos, déficit de vivienda y proyecciones futuras de estas variables, para poder trabajar sobre una base real en la programación urbana. Esto en la actualidad es bastante difícil dado que solo se cuenta con los Censos Nacionales de Población y Vivienda que en el Ecuador se realizan cada 12 años; pasando esta información a ser obsoleta en un período de tiempo muy corto, por este motivo creo conveniente aprovechar los datos censales para construir un Marco de Muestreo y obtener una muestra para estimar, en base a encuestas en hogares, los siguientes aspectos:

1.- Incremento poblacional:

Número de personas por hogar, por edad y sexo.

2.- Condiciones de vida de la población de acuerdo

a:

- Vivienda. (Número de hogares por vivienda, - metros cuadrados por habitante, etc.)
- Salario. (ingreso por persona u hogar)
- Gastos. (Alimentación, vivienda, transporte educación, luz, agua, etc.)
- Ocupación o actividad (Profesionales, funcionarios administrativos, comerciantes, - servicios, agricultores, obreros, campesinos, etc.)
- Grado de instrucción. (Superior, Secundaria, primaria, analfabeto.)
- Sexo, estado civil, etc.
- Otros. (Control de natalidad, consumo de alcohol, drogas, etc.)

3.- Migraciones. (Origen y trayectoria de la población.

Es necesario aclarar que por encontrarme realizando mis estudios de posgrado en la ciudad de México, me resulta muy difícil obtener todos los datos necesarios para obtener la muestra requerida, por lo tanto el objetivo en este caso es tener todo el diseño de muestra planeado, con los elementos de juicio de que dispongo - acá, para una vez que regrese a mi país pueda aplicar - dicho diseño a una investigación de ésta naturaleza.

La intención de este trabajo es adquirir la experiencia necesaria para apoyar a los procesos de planificación e investigación.

1 DESCRIPCION DE LA ZONA DE ESTUDIO:

Lo que en adelante denominaré el "Área Metropolitana de Cuenca" está formada por la ciudad de Cuenca (área urbana) y el área periférica (área rural y una serie de núcleos periféricos que se encuentran estrechamente ligados a la ciudad). La determinación de la superficie que abarca dicha "Área Metropolitana" y la definición de las áreas básicas de planificación, se hizo básicamente en función de:

- a) Características topográficas.
- b) Niveles de concentración poblacional.
- c) Sistema vial. (Ejes de desarrollo)
- d) Sectores Censales.

Se tomaron en cuenta estas características para de limitar geográficamente cada área de planificación. Las características topográficas en la medida que existen accidentes geográficos que hacen difícil la expansión urbana en ciertas partes del territorio, así como su accesibilidad a la ciudad de Cuenca que es el núcleo urbano mayor.

Se consideraron también los niveles de concentración poblacional para definir los núcleos periféricos y su área de influencia.

El sistema vial sirvió también para la determinación de las áreas básicas de planificación, ya que una vía de comunicación generalmente determina concentración poblacional alrededor de ella y facilita el transporte de bienes y personas hacia la ciudad.

Finalmente se consideran también los sectores censales que permiten aprovechar una serie de datos obtenidos en los censos de Población y vivienda para la definición de las áreas de Planificación de lo que se considerara como "Área Metropolitana de Cuenca" que son las siguientes:

Áreas de Planificación	Superficie	Población
1.- Cuenca	4.140 km ²	111018 hab.
2.- Checa Chiquintad	1.205 "	3537 "
3.- Ricsurte	1.802 "	7625 "
4.- Llacao	3.170 "	2598 "
5.- Paccha-Nulti	9.181 "	4179 "
6.- Turi	1.623 "	3340 "
7.- Tarqui	8.696 "	3064 "
8.- Baños	2.402 "	8550 "
9.- San Joaquín	828 "	3482 "
10.- Sayausí	1.242 "	2728 "
11.- Sinincay	3.567 "	7231 "
12.- El Valle	1.000 "	8157 "
TOTAL AREA METROPOLITANA	38.456 "	165508 "

Tomado de: Diagnóstico para el "Área Metropolitana de Cuenca". Elaborado por CONSULPLAN.

El área urbana de Cuenca se divide en 14 Zonas censales, con sus respectivos sectores y manzanas.

La población del área periférica está distribuida en concentrada y dispersa de la siguiente manera:

<u>Area de Planificación</u>	<u>Población (N^o de hab.)</u>		
	<u>Concentrada</u>	<u>Dispersa</u>	<u>Total</u>
Checa-Chiquintad	934	2603	3537
Ricaurte	897	6728	7625
Llacao	543	2055	2598
Paccha-Multi	523	3656	4179
Turi	453	2887	3340
Tarqui	169	2895	3064
Baños	1643	6907	8550
San Joaquín	854	2628	3482
Saysusi	688	2040	2728
Sinincay	346	6885	7231
El Valle	<u>254</u>	<u>7902</u>	<u>8156</u>
TOTAL AREA PERIPERICA	7304	47186	54490

Tomado de: Diagnóstico del "Área Metropolitana de Cuenca. Elaborado por: CONSULPLAN.

El área periférica está dividida en sectores de acuerdo a la superficie y población de cada área de planificación. Es necesario señalar que en las áreas de población concentrada existen manzanas, no siendo así en las áreas de población dispersa.

Se dispone en las Oficinas del Instituto Nacional de Estadísticas y Censos del Plano del Area Metropolitana de Cuenca, dividido en zonas y sectores censales y de un cuaderno cartográfico de los sectores censales, en donde se describe el número de manzanas, número de viviendas y su localización dentro de cada sector, además se dispone de la población censal de cada una de estas unidades.

De acuerdo a la información de que dispongo en la actualidad podían considerarse las siguientes características de las áreas a encuestarse; las mismas que son aproximaciones para fines de este trabajo.

1) Ciudad de Cuenca

Población (Censo 1974)	111018 habitantes
Superficie	4.14 km ²
Sectores censales	36
Población por sector	de 2 a 10 mil hab.
Número de manzanas por sector	de 20 a 50
Viviendas por manzana	de 20 a 50

2) Area Periférica

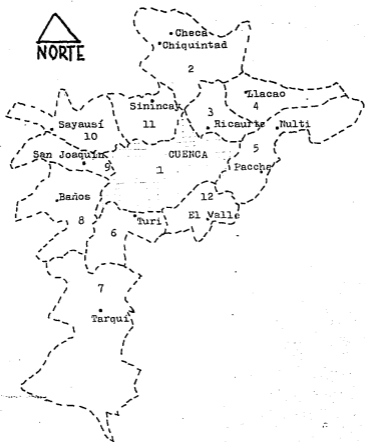
Población de cada área*	de 2500 a 6000 hab.
Población concentrada en c/área*	de 200 a 2000 hab.
Población dispersa den c/área*	de 2000 a 8000 hab.
Sectores censales	de 4 a 8 por área
Viviendas por manzana (pob concentrada)	de 15 a 30

* Se refiere a áreas de planificación

CROQUIS DEL AREA METROPOLITANA DE CUENCA

1 Area urbana

2-12 Área Periférica.



2 NIVELES DE GENERACION DE ESTADISTICAS:

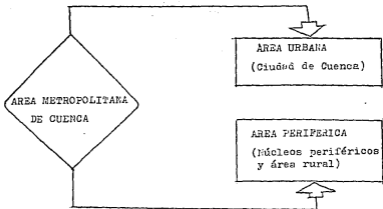
Dadas las características socio-económicas del "Área Metropolitana de Cuenca" es necesario producir información a nivel de toda ella, y para cada una de las dos áreas que se definen a continuación:

— La una formada por la ciudad de Cuenca, que se caracteriza por tener una población netamente urbana, un amanzamiento regular, estar directamente administrada por el "Municipio", el mismo que ejerce un control sobre todo lo que son nuevas edificaciones, croquis de las propiedades, enajenación o adquisición de viviendas, etc.

— La otra estaría formada por todas las áreas periféricas de planificación que rodean a la ciudad, cada una de ellas con un centro de población concentrada, en donde existe un amanzamiento, y una población dispersa - dedicada a la agricultura o la ganadería según el caso. Estas áreas son muy importantes para la ciudad en cuanto a que son proveedores de la mayor cantidad de artículos agrícolas y artesanales de consumo inmediato para la población urbana.

Esta división obedece a la necesidad de obtener inferencias por separado para cada una de estas dos áreas que presentan características diferentes para efectos de planificación e investigación.

Estas dos áreas están claramente definidas, tanto geográficamente como en los documentos censales, los mismos que servirán de base a la selección y determinación de la muestra.



3 DEFINICION DE UNIDADES DE MUESTREO Y ESTRATIFICACION:

Como no se dispone en la actualidad de los datos necesarios, no se determinará el tamaño de muestra, esto se hará posteriormente, cuando se vaya a aplicar la encuesta, de acuerdo a la variabilidad de la característica a estudiar y la disponibilidad de recursos.

Dentro del área urbana, o sea en la ciudad de Cuenca se definen las siguientes unidades de muestreo:

Unidades Primarias:

Se considerarán como unidades primarias las manzanas, que son claramente identificables y se tiene el ligto de las mismas según Censo.

Unidades Secundarias:

Se considerarán como unidades secundarias los hogares, entendiéndose como tal a : "una persona que vive sola, o dos o más personas que unidas o no por relaciones de parentesco, ocupan habitualmente una misma vivienda u otra clase de alojamiento o parte de ella y consumen en común alimentos y otros bienes indispensables para la vida, mediante un fondo general aportado por los miembros que perciben ingresos".*

*Encuesta Nacional de Ingresos y Gastos. Mexico 1977. Secretaría de Programación y Presupuesto.

Las unidades de tercera etapa dentro de la Ciudad de Cuenca serán los individuos, se generarán estadísticas a este nivel cuando se requiera estimar características poblacionales a nivel individual.

En el Área Periférica, por tener características diferentes se definen las siguientes unidades de muestreo:

Unidades Primarias:

Por la dispersión de la población en ésta área se tomarán como unidades primarias de muestreo los sectores censales, de esta manera lograríamos una muestra más concentrada y por lo tanto más accesible.

Unidades Secundarias:

Las unidades secundarias serían entonces, las manzanas censales, dentro de los núcleos periféricos (Población concentrada), o las áreas equivalentes, en la zona de población dispersa.

En el área periférica se consideran también unidades de tercera y cuarta etapa, según se quieran estimar características de la población a nivel de hogares o de individuos respectivamente.

Estratificación:

Antes de determinar como se estratificarían las manzanas en el área urbana y los sectores en el área periférica del "Área Metropolitana de Cuenca", para los fines de esta encuesta, mencionaré, en qué consiste y cuando se usa este esquema de muestreo.

El Muestreo Estratificado divide la población a muestrearse en subpoblaciones distintas llamadas estratos. Dentro de cada estrato se selecciona una muestra separada a partir de las unidades de muestreo de ese estrato. De la muestra obtenida en cada estrato se calculan estimadores: medias, varianzas, etc. Estos estimadores por estrato se ponderan adecuadamente para formar una estimación combinada de la población total.

Se usa la estratificación principalmente por las siguientes razones:

1) Para disminuir las varianzas de los estimadores. La Varianza disminuye según el grado en que difieran las medias por estrato y el grado de homogeneidad existente dentro de los estratos.

2) Se pueden formar los estratos para usar diferentes métodos de selección y procedimientos de estimación

en cada uno de ellos. Esto es necesario cuando la distribución física de algunas partes de la población difiere radicalmente. Cuando hay diferencias en las listas disponibles para diferentes partes de la población, y cuando la naturaleza diversa de los elementos de la población requiera del uso de diferentes procedimientos.

3) Los estratos pueden establecerse porque las subpoblaciones dentro de ellos también se definen como dominio o sea una parte de la población para la que se ha planteado estimaciones separadas en el diseño de la muestra en estudio.

4) Cuando la población está formada por unidades heterogéneas y tenemos una idea previa de los grupos de unidades homogéneas entre sí, es conveniente agrupar estas unidades y formar estratos.

Los estratos agrupan unidades homogéneas dentro de ellos y son en cambio heterogéneos entre ellos. Mientras mayor sea la homogeneidad dentro de los estratos más se puede reducir la variabilidad de estimadores que no sean de razón y regresión, sino habrá que estratificar para reducir la variabilidad de ellos.

Estratificación dentro del "Área Metropolitana de Cuenca":

Dentro de las dos áreas determinadas, se podrían formar a su vez estratos con unidades homogéneas entre sí. En el área urbana de la ciudad de Cuenca se estratifican manzanas de acuerdo a las características socio-económi-

cas de la población que reside en ella, y de acuerdo al destino principal que se da a las edificaciones en cuanto a su uso.

En el área periférica (parte rural) se estratificarán los sectores censales de acuerdo a sus características socio-económicas, las mismas que por su relación dependiente de la ciudad de Cuenca, se agrupan de acuerdo al tipo de bienes y servicios que prestan a la misma.

Para una mayor precisión en la estratificación del área urbana creo conveniente no considerar las zonas censales, sino hacerlo más bien por manzanas, ya que es una división más detallada y se puede obtener mayor precisión, en cambio en una zona a veces se encuentran hogares que pertenecen a diferentes estratos.

La ciudad de Cuenca se dividirá entonces en los siguientes estratos:

- 1.- Estrato Comercial e Industrial.
- 2.- Estrato Residencial de capas altas y medias de altos ingresos.
- 3.- Estrato residencial de capas medias.
- 4.- Estrato residencial de capas bajas.
- 5.- Estrato en transición.

Cada uno de estos estratos podríamos definirlos así:

Comercial e Industrial: Está formado por la parte central de la ciudad de Cuenca que es la zona comercial, y los alre

dedores de la vía que comunica a Cuenca con la parte Nor~~te~~ del país y con la Costa, especialmente con las ciudades de Quito, que es la capital, y Guayaquil que es el puerto principal.

Estas zonas por sus características especiales concentran la actividad comercial, financiera e industrial de la ciudad, al mismo tiempo se asienta una buena parte de la población, que habita al interior de las manzanas, muchas veces en forma tugurizada por el alto costo de la renta en dicho sector. En la zona industrial se ubica otra parte de la población para localizarse cerca de su lugar de trabajo, son generalmente obreros de las fábricas que se ubican en esta zona y que por tener mayor accesibilidad prefieren localizarse cerca de sus fuentes de trabajo y sus viviendas son de tipo económico.

Residencial de capas altas y capas medias: La población de esta zona está formada por los sectores dominantes de la ciudad, generalmente son viviendas unifamiliares que se caracterizan por tener una buena porción de terreno - parte jardines, locales para coches y otras comodidades - propias de esta clase. Se ubican en las cercanías de la ciudad, disfrutando al mismo tiempo de las mejores condiciones de la infraestructura de la ciudad y de las grandes áreas que pueden conseguir en estos barrios para la construcción de sus residencias.

Residencial de capas medias: La formación de estos barrios es reciente, responde en la mayor parte de los casos a la

construcción por parte de organismos del Estado, de viviendas de interés social, en donde se ubican las capas medias que por sus ingresos no pueden adquirir residencias exclusivas, pero que disponen por lo menos de los recursos necesarios para volverse propietarios de este tipo de viviendas que son financiadas a largo plazo.

Residencial de Casas Bajas: Son barrios que se van formando en las afueras de la ciudad, su principal actividad es la artesanía, (herrereros, zapateros, cerámicos, sastres) y los servicios que prestan en la ciudad (lavanderas, plomeros, servidoras domésticas, jardineros, etc.). Muchas de las viviendas de este sector son autoconstruidas y generalmente carecen de los servicios sanitarios indispensables.

Estrato en transición: Estaría formado por las zonas que todavía no están totalmente urbanizadas, por lo tanto requieren un amanzanamiento especial, además no existe una definición clara de las características socio-económicas de su población, ya que existen tanto residencias de lujo como modestas viviendas.

El área periférica se estratificaría uniendo áreas de planificación, porque en este caso no se da la superposición entre características socio-económicas de la población y de los sectores censales. Se formarían entonces tres estratos que son:

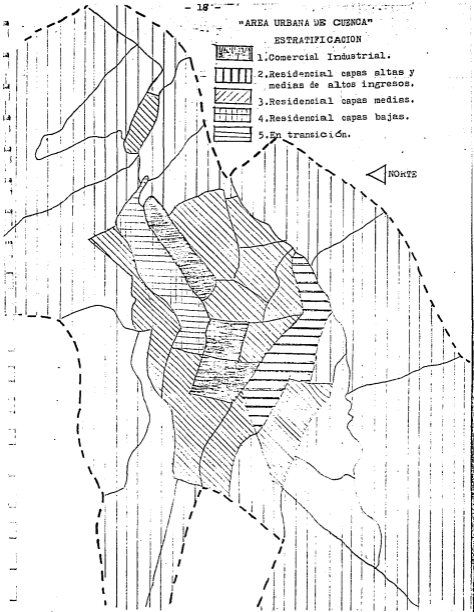
- 1) Estrato agrícola Artesanal.
- 2) Estrato Ganadero.
- 3) Estrato obrero artesanal.

"AREA URBANA DE CUENCA"

ESTRATIFICACION



- 1. Comercial Industrial.
- 2. Residencial capas altas y medias de altos ingresos.
- 3. Residencial capas medias.
- 4. Residencial capas bajas.
- 5. En transición.



Las características más importantes de cada uno de estos estratos son las siguientes:

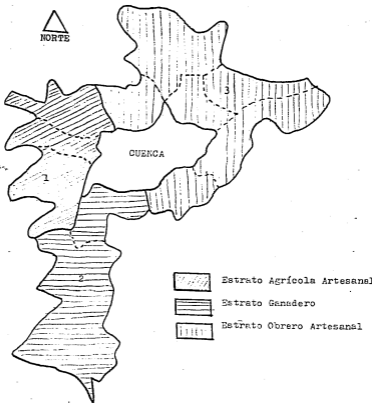
Agrícola Artesanal: Este estrato estaría formado por las áreas de planificación de San Joaquín, Sayausí y Baños, las mismas que combinan la producción agrícola a pequeña escala (maíz, verduras, cereales) con la artesanal (cannastos, bordados, muebles, etc.), todo esto dirigido al mercado de la ciudad de Cuenca y a otras áreas urbanas del país, especialmente Guayaquil.

Ganadero: Esta zona está un poco alejada del centro de la ciudad, está formada por las áreas de planificación de Turi y Tarquí, las mismas que por las características de sus tierras tienen buenos pastos y por lo tanto propicios para la ganadería; producen los lácteos para la ciudad de Cuenca y en pequeña escala para el resto del país. Su población está formada en su mayoría por campesinos que trabajan ya sea en las haciendas o en la ciudad como asalariados y habitan en pequeñas viviendas de tipo rural.

Obrero Artesanal: Las demás áreas de planificación periféricas se agrupan dentro de este estrato. Su población se dedica fundamentalmente a prestar servicios en la ciudad (albañiles, plomeros, carpinteros, etc.), combinando con actividades de tipo agrícola al poseer pequeñas parcelas de tierra que les permite cultivar productos como el maíz y tener algunos animales domésticos para su auto

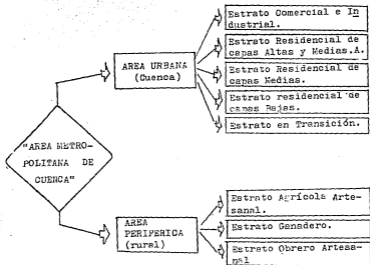
" AREA PERIPERICA "

ESTRATIFICACION



consumo o para la venta en los mercados de la ciudad, lo que les sirve para completar sus ingresos.

ESQUEMA DE ESTRATIFICACION



4 ESQUEMA DE MUESTREO:

Siendo las manzanas censales, en el área urbana, las unidades primarias de selección, ésta se haría con Muestre o Proporcional al Tamaño con Reemplazo, de acuerdo, al número de viviendas según censo, existentes en cada manzana. - Se usará este tipo de selección porque existen diferencias en el tamaño de cada manzana (número de viviendas), es útil controlar estas diferencias de tamaño porque pueden presentarse deficiencias en el trabajo de campo; el costo de la recolección de datos sería menor, ya que existe mayor probabilidad de selección, de manzanas con mayor número de viviendas; existiría una mayor eficiencia estadística - porque estarían más representadas las manzanas con mayor cantidad de viviendas; se evitaría tener manzanas sobrerrepresentadas y subrepresentadas según sea el caso.

La forma para extraer la muestra en este esquema es - la siguiente:

1.- Se obtienen los valores de X_1 (siendo X una variable auxiliar conocida), acumulados en un orden arbitrario.

2.- Se obtiene un número aleatorio entre 1 y X_1 .

X es el total conocido para los valores de X_1 .

3.- Se ve a que unidad corresponde el número extraído en el punto 2 y esa unidad se selecciona en la muestra.

4.- Se repite el proceso a partir del punto 2 hasta completar la muestra de tamaño n .

Unidad en la población	Tamaño	Tamaño acumulado
u_1	X_1	X_1
u_2	X_2	X_1+X_2
.
.
u_i	X_i	$X_1+X_2+\dots+X_i$
.
.
u_{N-1}	X_{N-1}	$X_1+X_2+\dots+X_{N-1}$
u_N	X_N	X
Total	X	

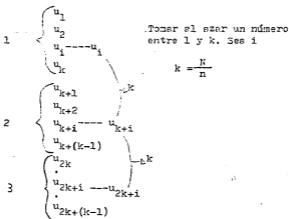
Tomado de: Conceptos muy elementales del Muestreo con énfasis en la determinación del tamaño de Muestra. I. Méndez R. IIMAS. UNAM.

Si el número aleatorio está entre $X_1 + X_2$ y $X_1+X_2+X_3$ se selecciona la unidad u_3 . Algunas unidades pueden aparecer más de una vez en la muestra.

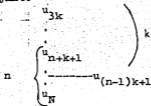
Dentro de cada Sector Censal tanto del área urbana de la ciudad como del área periférica, se dispone del listado de manzanas con el número de viviendas (datos del Censo Nacional de Población y Vivienda, complementados con los datos municipales de permisos de construcción y demolición de viviendas, lo que será complementado con un trabajo de campo en las manzanas seleccionadas), una vez seleccionadas las manzanas se hará un listado actualizado de hogares y se seleccionará los mismos con un muestreo sistemático ya que el arreglo de hogares no presenta ninguna regularidad y este esquema de muestreo sería equivalente a un muestreo aleatorio irrestricto, pero mucho más simple de llevarse a cabo ya que lo único que debería determinarse es un arranque aleatorio para la selección.

La forma de extraer la muestra en este esquema es la siguiente:

Conjunto



Conjunto



Los elementos de la muestra son: $u_i, u_{k+i}, u_{2k+i}, \dots, u_{(n-1)k+i}$.

Tomado de: Conceptos muy elementales del Muestreo con énfasis en la determinación del tamaño de la Muestra. Ignacio Méndez. IIMAS. UNAM.

5 ESTIMADORES:

Según el esquema planteado en este trabajo, los esti madores para el área urbana deberían construirse de dos e tapas, para estimar ciertas características de los hogares y de tres etapas si las estimaciones se hacen sobre varia- bles medidas en los individuos.

Para el área periférica tendríamos estimadores de 3 etapas para general información a nivel hogar, y de cuatro etapas para generar información a nivel de individuo. Sie n do estos bastante complejos, por lo que no los consideraré en el presente trabajo.

Para efectos de este trabajo se construirán únicamen- te los estimadores para el área urbana de la ciudad de Cuen- ca; tanto a nivel de estimación de características de los hogares, como a nivel de estimación de características de los individuos. Posteriormente determinaré los estimadores para el área periférica y el "Área Metropolitana de Cuenca".

Como se estratifican manzanas y se seleccionan con Pro- babilidad Proporcional al Tamaño (Número de viviendas según censo o conteos) con Reemplazo, y dentro de cada manzana se leccionada se seleccionan hogares con muestreo sistemático (en este caso equivalente a muestreo aleatorio irrestricto, por no existir un orden), los estimadores para generar in- formación a nivel hogar se construirían de la siguiente ma- nera:

Notación:

N = Total de manzanas en la ciudad de Cuenca.

N_h = Total de manzanas en el estrato H.

n_h = Total de manzanas en muestra del estrato H.

M = Total de hogares en la ciudad de Cuenca.

M_h = Total de hogares en el estrato H.

M_{hi} = Total de hogares en la manzana i del estrato H.

m_{hi} = Total de hogares en muestra de la manzana i del estrato H.

Y = Total de la característica en la ciudad de Cuenca.

Y_h = Total de la característica en el estrato H.

Y_{hi} = Total de la características en la manzana i del estrato H.

Y_{hij} = Valor de la característica en el j-ésimo hogar de la manzana i-ésima del estrato h-ésimo.

X = Total de la variable auxiliar en la ciudad de Cuenca. (número de viviendas según censo)

X_h = Total de la variable auxiliar en el estrato H.

X_{hi} = Total de la variable auxiliar en la manzana i del estrato H.

Estimador de un Total:

$$\hat{Y} = \sum_{h=1}^L \hat{Y}_h \quad h = 1, 2, \dots, L.$$

$$L = 5$$

$$\hat{Y}_h = \frac{X_h}{n_h} \sum_{i=1}^{n_h} \frac{M_{hi} \sum_{j=1}^{m_{hi}} Y_{hij}}{X_{hi}} = \frac{X_h}{n_h} \sum_{i=1}^{n_h} \frac{\hat{Y}_{hi}}{X_{hi}}$$

$$\hat{Y} = \sum_{h=1}^L \frac{X_h}{n_h} \sum_{i=1}^{n_h} \frac{\hat{Y}_{hi}}{X_{hi}}$$

Varianza del Estimador de un Total:

$$v(\hat{Y}) = \sum_{h=1}^L v(\hat{Y}_h)$$

$$v(\hat{Y}_h) = \frac{X_h \sum_{i=1}^{n_h} X_{hi} \left(\frac{Y_{hi}}{X_{hi}} - \frac{Y_h}{X_h} \right)^2}{n_h} + \frac{1}{n_h} \sum_{i=1}^{n_h} \left[k_{hi}^2 \left(1 - \frac{m_{hi}}{k_{hi}} \right) \frac{S_{whi}^2}{m_{hi}} \right] \frac{X_h}{X_{hi}}$$

en donde $S_{whi}^2 = \sum_{j=1}^{k_{hi}} \frac{(Y_{hij} - \bar{Y}_{hi})^2}{k_{hi}}$

Varianza Estimada del Estimador de un Total:

$$\hat{v}(\hat{Y}_h) = \frac{X_h^2}{n_h(n_h-1)} \sum_{i=1}^{n_h} \left(\frac{\hat{Y}_{hi}}{X_{hi}} - \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{\hat{Y}_{hi}}{X_{hi}} \right)^2 \quad \text{Estimador insesgado}$$

— Estimador de una Media: (o proporción cuando la variable toma valores 0,1).

$$\hat{Y}_c = \frac{\hat{Y}}{\hat{M}} = \frac{\sum_{h=1}^L \hat{Y}_h}{\sum_{h=1}^L \hat{M}_h} = \frac{\sum_{h=1}^L \frac{X_h}{n_h} \sum_{i=1}^{n_h} \hat{Y}_{hi}}{\sum_{h=1}^L \frac{X_h}{n_h} \sum_{i=1}^{n_h} \frac{M_{hi}}{X_{hi}}} \quad \text{Estimador sesgado}$$

Como este es un estimador sesgado, la variabilidad al rededor del parámetro se mide con el Error Cuadrático Medio:

$$\begin{aligned} \text{ECM}(\hat{Y}) &= \frac{1}{X^2} \left[\sum_{h=1}^L \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{X_{hi}}{X_h} (Y_{hi} - \bar{Y}_c M_{hi} - Y_h + \bar{Y} M_h) \right. \\ &\quad \left. + \sum_{h=1}^L \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{1}{X_{hi}} M_{hi}^2 \left(1 - \frac{M_{hi}}{X_{hi}}\right) \frac{S_{hui}^2}{M_{hi}} \right] \end{aligned}$$

en donde:

$$S_{hui}^2 = \frac{1}{K_{hi} - 1} \sum_{j=1}^{K_{hi}} (Y_{hij} - \bar{Y}_{hi})^2$$

El Estimador del Error Cuadrático Medio estaría dado por:

$$\begin{aligned} \widehat{\text{ECM}}(\hat{Y}) &= \frac{1}{\hat{X}^2} \sum_{h=1}^L \frac{1}{n_h(n_h - 1)} \sum_{i=1}^{n_h} \left[\frac{K_{hi} (\hat{Y}_{hi} - \hat{Y})}{\frac{X_{hi}}{X_h}} \right. \\ &\quad \left. - \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{M_{hi} (\hat{Y}_{hi} - \hat{Y})}{\frac{X_{hi}}{X_h}} \right]^2 \end{aligned}$$

$$\widehat{ECM}(\widehat{Y}) = \frac{1}{M^2} \sum_{h=1}^L \frac{1}{n_h(n_h-1)} \sum_{i=1}^{n_h} \left[\frac{\widehat{Y}_{hi} - \widehat{Y}_{M_{hi}}}{\frac{X_{hi}}{X_h}} - \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{(\widehat{Y}_{hi} - \widehat{Y}_{M_{hi}})}{\frac{X_{hi}}{X_h}} \right]^2$$

Supongamos ahora que la unidad elemental es el individuo, entonces tenemos 3 etapas de selección: La primera de manzanas con Probabilidad Proporcional al Tamaño (viviendas según censo), la segunda de hogares con muestreo sistemático ; y la tercera los individuos dentro del hogar. (un censo de todos los individuos del hogar).

Usando la misma notación anterior, nos quedaría por incluir las siguientes:

Y_{hijk} = Valor de la característica en el individuo k del hogar j de la manzana i del estrato h.

L = Total de individuos en Cuenca.

L_h = Total de individuos en el estrato h.

L_{hi} = Total de individuos en la manzana i del estrato h.

L_{hij} = Total de individuos en el hogar j de la manzana i del estrato h.

- Estimador de un Total:

$$\widehat{Y} = \sum_{h=1}^L \widehat{Y}_h$$

En donde:

$$Y_h = \frac{X_h}{n_h} \sum_{i=1}^{n_h} \frac{\hat{Y}_{hi}}{X_{hi}}$$

En donde:

$$\hat{Y}_{hi} = K_{hi} \frac{\sum_{j=1}^{m_{hi}} Y_{hij}}{m_{hi}} = N_{hi} \hat{Y}_{hij}$$

En donde:

$$Y_{hij} = \frac{\sum_{k=1}^{L_{hij}} Y_{hijk}}{L_{hij}}$$

Varianza del Estimator de un Total:

$$V(\hat{Y}) = \sum_{h=1}^L v(\hat{Y}_h)$$

$$v(\hat{Y}_h) = \frac{1}{n_h} \sum_{i=1}^{N_h} \frac{X_{hi}}{X_h} \left(\frac{Y_i}{\frac{X_{hi}}{X_h}} - Y_h \right)^2 + \frac{1}{n_h} \sum_{i=1}^{N_h} \frac{K_{hi}(1 - \frac{m_{hi}}{m_{hi}}) \frac{S_{whi}^2}{m_{hi}}}{\frac{X_{hi}}{X_h}}$$

Como dentro del hogar se realiza un censo de todos los individuos que viven en él, la Varianza del estimador del Total sería igual que en el caso anterior. Luego entonces el estimador de la Varianza del Total es el mismo que en dos etapas.

Estimador de una Media:

$$\hat{Y} = \frac{\sum_{h=1}^L \hat{Y}_h}{\sum_{h=1}^L \hat{L}_h} = \frac{\sum_{h=1}^L \frac{X_h}{n_h} \sum_{i=1}^{n_h} \frac{K_{hi} \sum_{j=1}^{m_{hi}} \sum_{k=1}^{L_{hij}} Y_{hijk}}{m_{hi}}}{\sum_{h=1}^L \frac{X_h}{n_h} \sum_{i=1}^{n_h} \frac{K_{hi} \sum_{j=1}^{m_{hi}} L_{hij}}{m_{hi}}}$$

El Error Cuadrático Medio del estimador de la media sería el mismo que en el caso anterior.

Estimador del Error Cuadrático Medio de la Media Estimada:

$$\widehat{ECM}(\widehat{Y}) = \frac{1}{\widehat{M}^2} \sum_{h=1}^L \frac{1}{n_h(n_h-1)} \sum_{i=1}^{nh} \left[\frac{\widehat{Y}_{hi} - \widehat{Y} \widehat{L}_{hi}}{\frac{x_{hi}}{x_h}} \right]^2$$

$$= \frac{1}{n_h} \sum_{i=1}^{nh} \frac{(\widehat{Y}_{hi} - \widehat{Y} \widehat{L}_{hi})^2}{\frac{x_{hi}}{x_h}}$$

6 CONCLUSIONES:

De la experiencia obtenida en la realización de este trabajo se desprende un aspecto muy importante a tener en cuenta cuando se elabora el diseño de una muestra, que es lograr la mayor sencillez en la construcción de estimadores y varianzas de los mismos, a fin de que, ésta no se vuelva compleja y por lo tanto dificulte la obtención y manejo de la información que se genere en la muestra.

Además la mayor sencillez implicaría tener en cuenta al definir etapas en la muestra, que estas no sean muy numerosas, procurando al mismo tiempo optimizar los resultados, sin complicar demasiado el diseño.

Considero que el diseño de muestra realizado en el presente trabajo va a ser de mucha utilidad dentro de la Investigación y Planificación Urbana, en la medida que generará información en períodos muy cortos que permitirá mayor agilidad en la detección de ciertos fenómenos urbanos y en la programación de las actividades de Planificación.

7 Definiciones:

Áreas básicas de Planificación:

Es una unidad territorial, que presenta, características homogéneas, en cuanto a las actividades que realiza la población y las características geográficas; tienen como fin servir de base a la planificación operativa, para la realización de programas y proyectos de planificación urbana.

Niveles de concentración poblacional:

Se define en base a la densidad de población por hectárea; se considera - área de población concentrada cuando existen más de 50 habitantes por hectárea, y área de población dispersa - cuando existen menos de 50 habitantes por hectárea.

Sectores Censales:

Son unidades territoriales definidas por el Instituto Nacional de Estadísticas y Censos para la realización del levantamiento y las encuestas, - están definidas de acuerdo a la capacidad operativa de los encuestadores y se registra la información censal a base de estas unidades.

8 BIBLIOGRAFIA:

- 1.-Ignacio Mendez. Conceptos muy elementales del Muestreo con énfasis en la determinación práctica del tamaño de la muestra. IIMAS.Serie Azul.UNAM.1976.
- 2.- Leslie Kish. Muestreo de Encuestas. Editorial Trillas.Mexico 1975
- 3.-William G. Cochran. Técnicas de Muestreo. Editorial CECSA. Mexico 1974.
- 4.-Des Raj. Sampling Theory. Editorial McGraw-Hill. USA.1968
- 5.- Alianza para el Progreso. ATLANTIDA. Para encuestas en Hogares
- 6.- Consulplan. Diagnóstico del "Area Metropolitana de Cuenca.
- 7.-Secretaría de P. y Presupuesto Marco de hogares. Mexico.
- 8.- Instituto Nal. de Estadística. "Muestra Maestra" España.
- 9.- Hansen, Hurwitz. Sample Survey. Methods and Theory.
- 10.-Secretaría de P.yPresupuesto Encuesta Nacional de Ingresos y Gastos. Instructivo para el llenado del Cuestionario. Mexico 1977.